

ABSTRAK

Kemajuan teknologi digital telah membawa perubahan signifikan pada cara masyarakat berbagi informasi dan pengalaman, terutama melalui aplikasi mobile. Salah satu aplikasi yang berperan penting bagi masyarakat dan wisatawan di Yogyakarta adalah JogjaKita. Aplikasi ini menyediakan berbagai fitur seperti panduan perjalanan, informasi publik, dan layanan sehari-hari. Namun, kualitas dan efektivitas layanan aplikasi masih menjadi perhatian yang perlu dievaluasi. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengguna terhadap aplikasi JogjaKita, menggunakan metode Lexicon Based dan Naive Bayes untuk mengklasifikasikan sentimen positif, negatif, dan netral.

Penelitian dilakukan dengan mengambil data ulasan dari Google Play Store menggunakan teknik web scraping. Data yang terkumpul melalui tahapan preprocessing teks, termasuk case folding, normalisasi, penghapusan kata tidak penting (stopword removal), dan stemming. Dua metode utama digunakan untuk analisis sentimen: metode Lexicon Based yang memanfaatkan kamus sentimen untuk menilai polaritas kata-kata dalam teks, serta Naive Bayes yang berbasis probabilistik untuk mengklasifikasikan data berdasarkan pola tertentu. Model dikembangkan menggunakan teknik pembobotan TF-IDF untuk meningkatkan akurasi prediksi.

Hasil penelitian menunjukkan bahwa metode Naive Bayes menghasilkan akurasi yang lebih tinggi dibandingkan *metode Lexicon Based*, menguatkan temuan penelitian sebelumnya. Evaluasi menggunakan metrik akurasi, presisi, dan recall menunjukkan keunggulan Naive Bayes, dengan akurasi mencapai lebih dari 90%. Sementara itu, metode Lexicon Based memiliki kekurangan dalam menangkap konteks kata-kata yang ambigu atau sarat makna ganda. Penelitian ini memberikan kontribusi signifikan dalam bidang analisis sentimen, terutama dalam membantu pengembang aplikasi untuk memahami kebutuhan dan harapan pengguna, serta meningkatkan kualitas layanan aplikasi JogjaKita.

Berdasarkan hasil penelitian dengan menggunakan 1733 data dengan perbandingan data latih dan data uji yaitu 80%:20%, dimana data latih sebanyak 1386 data dan data uji sebanyak 347 data, dengan pengujian menggunakan tabel *confusion matrix* dihasilkan nilai akurasi sebesar 81%.

Kata kunci: Analisis Sentimen, Naive Bayes, *Lexicon Based*, JogjaKita.

ABSTRACT

The advancement of digital technology has brought significant changes to the way people share information and experiences, especially through mobile applications. One of the applications that plays an important role for the community and tourists in Yogyakarta is JogjaKita. This application provides various features such as travel guides, public information, and daily services. However, the quality and effectiveness of the application's services remain issues that need to be evaluated. This study aims to analyze user sentiment towards the JogjaKita application, using the Lexicon Based and Naive Bayes methods to classify positive, negative, and neutral sentiments.

The research was conducted by collecting review data from Google Play Store using web scraping techniques. The collected data underwent text preprocessing stages, including case folding, normalization, stopword removal, and stemming. Two main methods were used for sentiment analysis: the Lexicon Based method, which utilizes a sentiment lexicon to assess the polarity of words in text, and the Naive Bayes method, which uses probabilistic modeling to classify data based on specific patterns. The model was developed using TF-IDF weighting techniques to enhance prediction accuracy.

The results showed that the Naive Bayes method achieved higher accuracy compared to the Lexicon Based method, reinforcing findings from previous studies. Evaluation metrics, including accuracy, precision, and recall, demonstrated the superiority of Naive Bayes, with accuracy exceeding 90%. Meanwhile, the Lexicon Based method had limitations in capturing the context of ambiguous or polysemous words. This study makes a significant contribution to the field of sentiment analysis, particularly in assisting application developers to understand user needs and expectations, as well as improving the quality of JogjaKita's services.

Based on the research results using 1,733 data points with a training and testing data ratio of 80%:20%, where the training data consisted of 1,386 entries and the testing data consisted of 347 entries, the evaluation using a confusion matrix yielded an accuracy score of 81%.

Keywords: *Sentiment Analysis, Naive Bayes, Lexicon Based, JogjaKita.*